GP-UCB proof sketch

Gaussian Process Optimization in the Bandit Setting: No Regret and Experimental Design

Niranjan Srinivas, Andreas Krause, Sham M. Kakade, Matthias Seeger

Chernoff bound and union bound

- Finite sampling set D
- $|f(x) - \mu_{t-1}(x)| \leq \beta_t^{\frac{1}{2}} \sigma_{t-1}(x)$ for all $x \in D$ and $t \geq 1$
- Proof:
  - Chernoff bound: for **each** $x \in D$ and $t \geq 1$
  - $\Pr\left(|f(x) - \mu_{t-1}(x)| > \beta_t^{\frac{1}{2}} \sigma_{t-1}(x)\right) \leq e^{-\frac{\beta_t}{2}}$
- By union bound:
  - $\Pr\left(|f(x) - \mu_{t-1}(x)| \leq \beta_t^{\frac{1}{2}} \sigma_{t-1}(x)\right) > 1 - |D|\, e^{-\frac{\beta_t}{2}}$

Chernoff bound and union bound

- Choose $|D|e^{-\frac{\beta_t}{2}} = \frac{\delta}{\pi_t}$ with $\sum_t \frac{1}{\pi_t} = 1$
- So by union bound on $t \geq 1$
  - $\Pr\left(|f(x) - \mu_{t-1}(x)| \leq \beta_t^{\frac{1}{2}} \sigma_{t-1}(x)\right) > 1 - \sum_t |D| e^{-\frac{\beta_t}{2}} = 1 - \delta$

Regret

- Lemma 5.2. Fix $t \geq 1$, if $|f(x) - \mu_{t-1}(x)| \leq \beta_t^{\frac{1}{2}} \sigma_{t-1}(x)$ for all $x \in D$, then the regret γ_t at time is bounded by $2\beta_t^{\frac{1}{2}} \sigma_{t-1}(x_t)$
- Proof:
  - By definition of x_t: $\mu_{t-1}(x_t) + \beta_t^{\frac{1}{2}} \sigma_{t-1}(x_t) \geq \mu_{t-1}(x^*) + \beta_t^{\frac{1}{2}} \sigma_{t-1}(x^*) \geq f(x^*)$
  - Therefore the regret:
    - $\gamma_t = f(x^*) - f(x_t) \leq \beta_t^{\frac{1}{2}} \sigma_{t-1}(x_t) + \mu_{t-1}(x_t) - f(x_t) \leq 2\beta_t^{\frac{1}{2}} \sigma_{t-1}(x_t)$

Information Gain

- Lemma 5.3:
  - The information gain for the points selected can be expressed in terms of the predictive variances. If $f_T = \left(f(x_t)\right) \in R^T$:
  - $I(y_T; f_T) = \frac{1}{2}\sum_{t=1}^{T} \log(1 + \sigma^{-2}\sigma_{t-1}^2(x_t))$
  - Proof: By induction,
    - $I(y_T; f_T) = H(y_T) - \frac{1}{2}\log|2\pi e\sigma^2 I|$
    - $H(y_T) = H(y_{T-1}) + H(y_T|y_{T-1}) = H(y_{T-1}) + \frac{\log\left(2\pi e\left(\sigma^2+\sigma_{t-1}^2(x_T)\right)\right)}{2}$

Regret Bound

- Lemma 5.4.
- Pick $\delta \in (0,1)$ and let β_t be defined as in Lemma 5.1. Then, the following holds with probability $\geq 1-\delta$
- $\sum_{t=1}^{T} r_t^2 \leq \beta_T C_1 I(y_T; f_T) \leq C_1 \beta_T \gamma_T \quad \forall\, T \geq 1$
- Proof:
  - By Lemma 5.2, $\gamma_t^2 \leq 2\beta_t \sigma_{t-1}^2(x_t)$
  - By Lemma 5.3, $I(y_T; f_T) = \frac{1}{2}\sum_{t=1}^{T} \log(1+\sigma^{-2}\sigma_{t-1}^2(x_t))$
    - Where $\log(1+\sigma^{-2}\sigma_{t-1}^2(x_t)) \sim C_0 \sigma_{t-1}^2(x_t)$

Regret Bound

- By Cauchy inequality, $(\sum_{t=1}^{T} \gamma_t)^2 \leq T \sum_{t=1}^{T} r_t^2 \leq T C_1 \beta_T \gamma_T$
- So the total regret $\sum_{t=1}^{T} \gamma_t \leq \sqrt{T C_1 \beta_T \gamma_T} = O(\sqrt{T}) O(\sqrt{\beta_T \gamma_T})$

In our case: Chernoff Bound

- Consider $f(x) = \sum_i g_i(x) f_i(x) \sim N\left(\mu_t(x), \sigma_t^2(x)\right)$
- Where $\mu_t(x) = \sum_i g_i(x)\mu_{i,t}(x)$, $\sigma_t^2(x) = \sum_i g_i^2(x)\sigma_{i,t}^2(x)$
- So it is fine to have the same result.

In our case: Regret

- $\gamma_t = f(x^*) - f(x_t) \leq \beta_t^{\frac{1}{2}} \sigma_{t-1}(x_t) + \mu_{t-1}(x_t) - f(x_t) \leq 2\beta_t^{\frac{1}{2}} \sigma_{t-1}(x_t)$
- $\sigma_{t-1}^2(x_t) = \sum_i^I g_i^2(x_t) \sigma_{i,t-1}^2(x_t)$

In our case: Information Gain

- Original: $I(y_T; f_T) = \frac{1}{2}\sum_{t=1}^{T} \log(1 + \sigma^{-2}\sigma_{t-1}^2(x_t))$
- Here: $I(y_T; f_T) = \frac{1}{2}\sum_{i=1}^{I}\sum_{t=1}^{T} \log(1 + \sigma^{-2}\sigma_{i,t-1}^2(x_t))$ which is a little different from directly substituting the $\sigma_{t-1}^2(x_t)$ by entire aggregated $\sum_i^I g_i^2(x_t)\sigma_{i,t-1}^2(x_t)$
- $I(y_T; f_T) = \frac{1}{2}\sum_{i=1}^{I}\sum_{t=1}^{T} \log(1 + \sigma^{-2}\sigma_{i,t-1}^2(x_t)) \geq \boldsymbol{\frac{1}{2}\sum_{i=1}^{I}\sum_{t=1}^{T} \sigma^{-2}\sigma_{i,t-1}^2(x_t)}$

In our case: Summary

- What we have
  - $I(y_T; f_T) = \frac{1}{2}\sum_{i=1}^{I}\sum_{t=1}^{T}\log(1+\sigma^{-2}\sigma_{i,t-1}^2(x_t)) \geq \boldsymbol{\frac{1}{2}\sum_{i=1}^{I}\sum_{t=1}^{T}\sigma^{-2}\sigma_{i,t-1}^2(x_t)}$
  - $\gamma_t = f(x^*) - f(x_t) \leq \beta_t^{\frac{1}{2}}\sigma_{t-1}(x_t) + \mu_{t-1}(x_t) - f(x_t) \leq 2\beta_t^{\frac{1}{2}}\sigma_{t-1}(x_t)$
    - Or equivalently, $\gamma_t^2 \leq 4\beta_t \sum_i^I g_i^2(x_t)\sigma_{i,t-1}^2(x_t)$
- We want to relate $\gamma_1 + \gamma_2 + \cdots + \gamma_T$ to the information gain $I(y_T; f_T)$

In our case: Next Step

- 1. Use some clever Cauchy inequality or something else to relate the total regret and the information gain.
- 2. Follow the remaining proof in GP-UCB paper to derive a bound on the maximum information gain.
  - They use relaxation and submodular optimization…
- 3. Back to our first step, learn the transition function (each $g_i(x)$)…